Vermont Mental Health Performance Indicator Project
DDMHS, Weeks Building, 103 South Main Street, Waterbury, VT 05671-1601 (802-241-2638)

MEMORANDUM

TO: Vermont Mental Health Performance Indicator Project
Advisory Group and Interested Parties

FROM: John Pandiani
Monica Simon

DATE: March 13, 2002

RE: Education Test Findings Presented at Tampa Children's MH Conference

We have enclosed a copy of the handout from our recent presentation on "Using Educational Test Scores to Evaluate Children's Services" at the 15th Annual Research Conference: A System of Care for Children's Mental Health. The presentation, which focused on gender and regional differences, in addition to methodological issues, was well received and stimulated a number of interesting suggestions for further analysis. This, our first national presentation of findings, followed our first receipt of educational test score data by less than 10 months.

Please feel free to share this handout (and any of our earlier reports) with anyone you think may be interested. We are particularly interested in learning about other projects that are using educational test scores (or other measures of school participation and/or performance) to evaluate mental health or other children's services programs. As always, we look forward to hearing from you at edtestresearch@ddmhs.state.vt.us.

If this e-mail was forwarded to you and you would like to be part of our regular distribution, just send an email to edtestresearch@ddmhs.state.vt.us.

USING EDUCATIONAL TEST SCORES TO EVALUATE CHILDREN'S SERVICES

John A. Pandiani, Ph.D.
jpandiani@ddmhs.state.vt.us
Chief of Research and Statistics

Monica M. Simon, M.S.
msimon@ddmhs.state.vt.us
Senior Research and Statistics Analyst

Vermont Department of Developmental and Mental Health Services
103 South Main Street, Waterbury, Vermont 05671-1610

Steven M. Banks, Ph.D.
bristob@together.net
Chief Statistician, The Bristol Observatory
521 Hewitt Road, Bristol, Vermont 05443

ABSTRACT

Standardized educational test results are becoming an important part of the children's services data infrastructure. This presentation reports the results of using these test scores to evaluate mental health programs. Results indicate that standardized test scores are related to an important non-school outcome indicator. School participation for CMH clients is lower than for other young people and there are significant differences among regional treatment programs. School performance for CMH clients is lower than for other students statewide, but in some regions of the state service recipients performed as well as other students. When pre-treatment test scores are compared to post-treatment scores, there are substantial differences between male and female service recipients.

Presented at the 15th Annual Research Conference
A System of Care for Children's Mental Health: Expanding the Research Base
March 3-6, 2002 - Tampa Hyatt Regency Westshore, Tampa, Florida

This project was supported, in part, by a Center for Mental Health Services Sixteen State Indicator Pilot Project Grant #5HR1SM52625-03. This report reflects the views and opinions of the authors. It does not necessarily reflect the official position or policy of CMHS, SAMHSA, or HHS.

We thank Michael Curtis, Ed.D. Children's Services Director, Washington County Mental Health Services and Vicki C. Hornus, Special Project Coordinator, Vermont Department of Education for helping us to see our findings from the perspective of people in different parts of Vermont's system of care.

Standardized educational test scores are becoming an increasingly important part of the children's services data infrastructure. The recently enacted Elementary and Secondary Education Act of 2001 requires statewide reading and mathematics tests each year in grades 3 through 8 by 2005-06. This legislation follows on efforts in many states to use standardized testing to introduce higher levels of accountability in education. The Vermont Department of Education implemented a statewide, standardized testing program, the New Standards Reference Exams, in 1998. Since that time, public school students throughout the state have been tested on an annual basis in Mathematics and English/Language at the fourth, eighth, and tenth grade levels. Vermont's standardized tests are not used to determine student advancement, but are used as a core measure of school performance.

This proliferation of standardized testing is occurring at a time when children's mental health programs and service systems are being increasingly called upon to demonstrate their favorable impact on school participation and performance [1,2]. Within this context, the Vermont Mental Health Performance Indicator Project (PIP) is exploring the relevance and utility of these test scores for evaluating community mental health programs. This presentation reports on the results of our preliminary analysis of standardized test scores from Vermont's first four years of statewide testing.

The Philosophy of the PIP

This project is strongly influenced by a philosophy of evaluation research that is data-based, incremental, strengths-based, and perspectivistic. Our research process is data-based in that it begins with factual knowledge and moves to next questions. It is based almost exclusively on existing data resources (data we have laying around the house) rather than embarking on time consuming and expensive data collection initiatives. This process is incremental in that it answers one small question at a time, carefully considers the results, and then moves on to another small question. As answers to small questions accumulate, they are compared and potential patterns are considered and discussed. This process strives to remain "strengths-based" in that it focuses on what the findings tell us, not on other important questions and issues that are not addressed. Methodologically, we strive to focus on the strengths of research methods, not their weaknesses in order to not to let "the perfect" be the enemy of "the good". Finally, this process is perspectivistic. A large diverse group of "information consumers" comment on and suggest interpretations of findings as they become available. In this process, participants come to recognize the fact that different people frequently interpret findings differently and that no data are relevant to only one question.

Methodology

Data

The findings reported here are based exclusively on analysis of anonymous extracts from four existing databases. Anonymous, person level, extracts from the Department of Education's Mathematics Skills Assessment and English Language Reading: Basic Understanding tests for 4th, 8th, and 10th grade students during 1998 through 2001 provide our measures of school participation and performance. The number of young people represented in these education data sets averaged 20,743 per year.

Anonymous, person level, extracts from the Vermont Mental Health Division's Monthly Service Report database provide basic information on all young people who received community mental health services during the same time periods. The number of young people in the relevant age groups who were represented in the community mental health data set each averaged 2,279 per year.

Anonymous event level extracts from the Vermont District Court database provided basic information on all individuals who were charged with a crime in Vermont during the study period. The number of young people in the relevant age groups who were represented in this data set averaged 1,599 per year.
Finally, the Vermont Department of Health Annual Report on Population and Housing provided age and gender specific population figures at the town and county level for discrete age and gender categories for the time periods under examination.

Method

School participation, for purposes of this project, is demonstrated by completing a specified standardized test. Failure to compete the test constitutes non-participation. School participation rates for mental health service recipients are determined by measuring the number of people represented in both the mental health and the school test data set and presenting the results as a percent of all service recipients. School participation rates for the larger population of all young people were obtained by dividing the total number of young people tested by the total population of the state for the specified age group and presenting the results as a percentage. In order to provide a fair measure of school participation for mental health services recipients, their participation rate is compared to the participation rate for the general population and expressed as a "school participation ratio", where "1" indicates no difference, a score less than "1" indicates a lower participation rate by service recipients, and a score of higher than "1" indicates that service recipients participate at a higher rate than other students.

Our measure of school performance is the proportion of students who score at or above the predetermined "standard" for the test under examination. School performance rates for mental health service recipients are determined by measuring the number of young people represented in the mental health data set who are also represented in a data set of "high" test scorers and presenting the result as a percent of all service recipients. In order to provide a fair measure of school performance for mental health service recipients, their performance is compared to the performance of other students and expressed as a "school performance ratio", where "1" indicates no difference, a score less than "1" indicates lower performance by service recipients, and a score greater than "1" indicates that service recipients perform at a higher level than other students.

For purposes of determining the relationship between these school-based measures and other, non-school-based, outcome measures, school participation/performance rates are compared to rates of criminal justice involvement for 16 year old students in each of three time periods. Criminal justice involvement, for this analysis, involves being charged with a crime in a Vermont District Court. Criminal justice involvement was measured for young people in each of three groups. These groups include young people who scored at or above standard, young people who scored below standard, and young people who were not tested (on grade level). For the two groups that participated in the test, the number of young people who were charged with a crime is the number who are represented in both the test score data set and the criminal justice data set. The rate of criminal justice involvement for young people who did not participate is determined by comparing the number of young people who did not participate to the number who were in the criminal justice data set but were not in either the high or the low performance category.

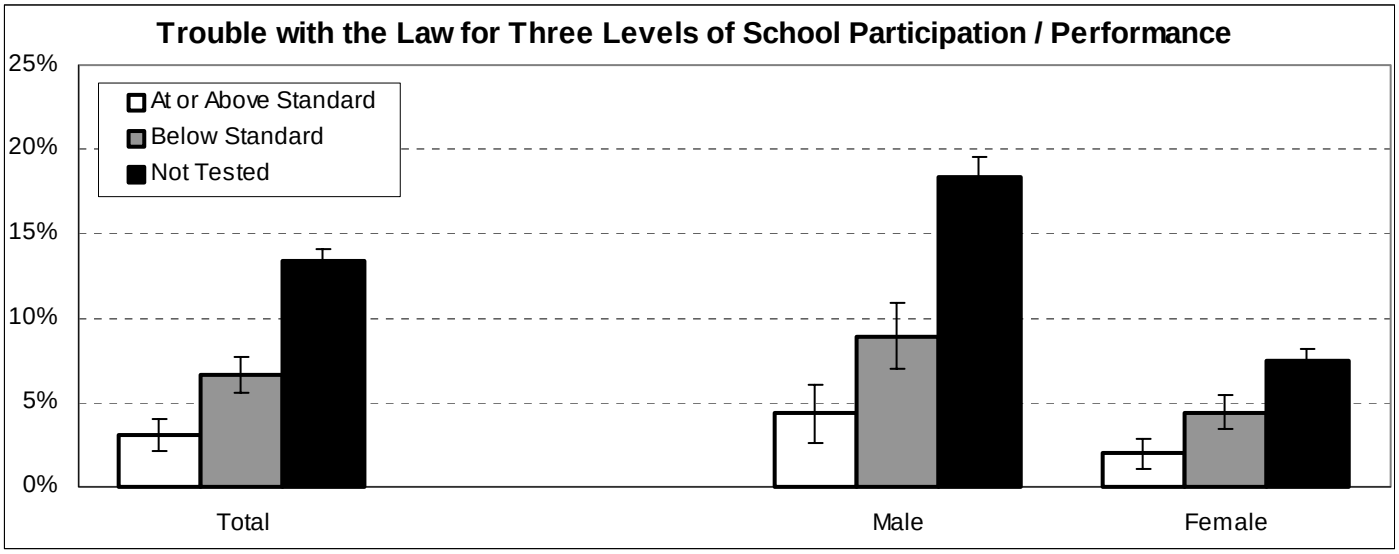
Because the data sets used in this analysis do not share unique person identifiers, Probabilistic Population Estimation was used to determine the number of people who appeared in combinations of data sets. Probabilistic Population Estimation is a statistical procedure that provides valid and reliable measures of the size and overlap of data sets that do not include unique person identifiers [3]. These estimates are based on comparisons of the distribution of dates of birth in the data sets to the known distribution of dates of birth in the general population. A technical description of the procedure and a list of published papers that describe and demonstrate the use of this procedure are provided at the end of this handout.

Findings

School Test Scores and Criminal Justice Involvement

There were statistically significant differences in the likelihood of criminal justice involvement between young people who scored at or above standard, those who scored below standard, and those who did not participate in the testing. Young people who scored at or above the standard were the least likely to be charged with a crime and young people who did not participate were the most likely to be charged with a crime. During the year after the test, almost 12% of the young people who did not participate were

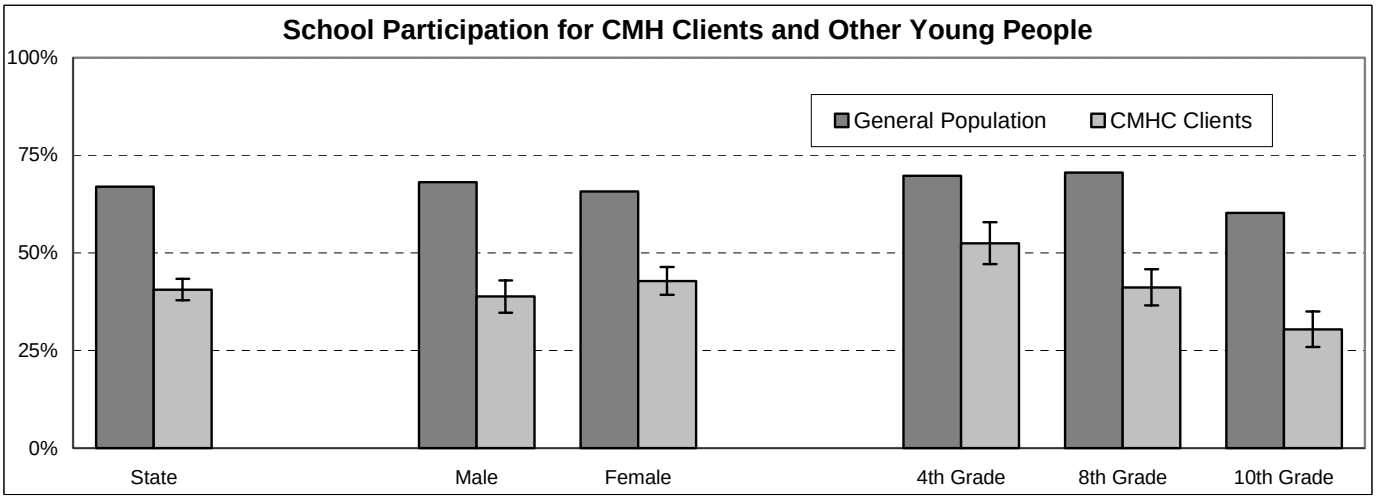
charged with a crime, compared to less than 4% of the young people who scored at or above the standard on the test. Boys were more likely than girls to be charged with a crime in all three groups, and the difference in rates of criminal justice involvement between groups was greater for boys than for girls.



When cumulative rates of criminal justice involvement after testing (first year vs. two years vs. three years) of low scorers are compared to the same rates for high scorers, the strength of the relationship between test scores and criminal justice involvement tends to decrease over time. Young people in the low test score group were 2.5 times more likely than high scorers to be charged with a crime during the test year but this elevated risk decreased to 1.9 at the end of three years. This decrease is much more substantial for young women (2.5 to 1.4).

School Participation

Recipients of community mental health services were less likely to take part in Vermont's statewide school testing than other students. Statewide, only 41% of the service recipients participated in the Mathematics Skills Assessment, compared to 67% of all students. Female service recipients were more likely to participate in school testing than male service recipients in the 4th and 8th grade, but this difference disappeared in the 10th grade. These patterns were evident for both Mathematics and English tests.



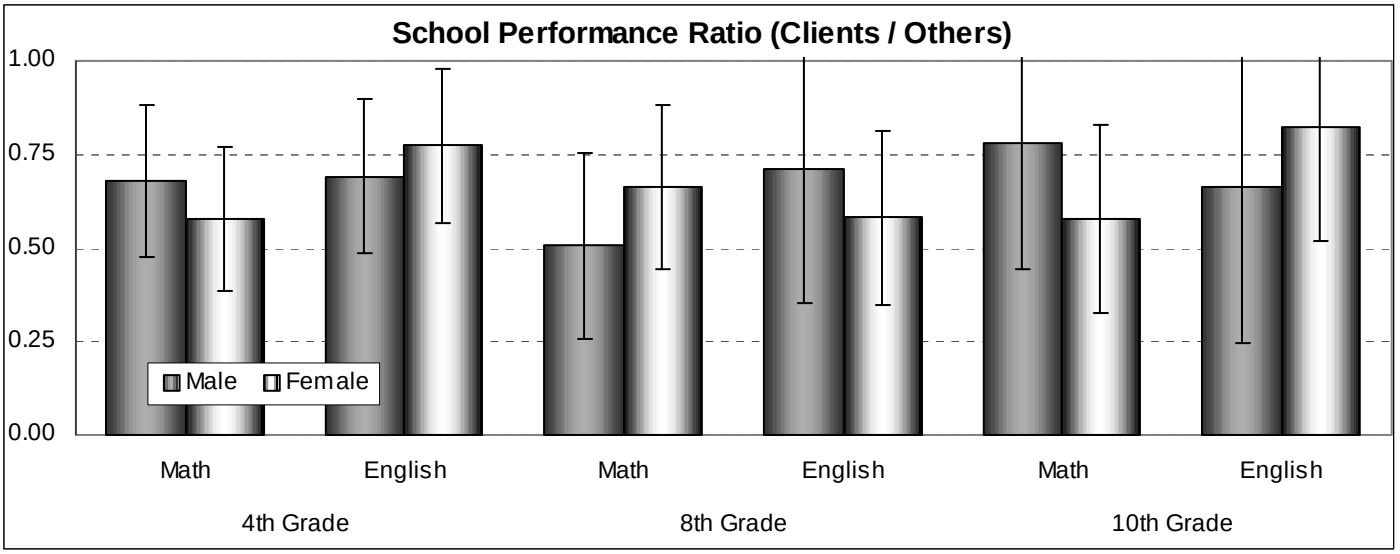
There were significant differences among Vermont's regional community mental health programs in relative test participation (controlling for general population participation rates). Young clients of Lamoille County Mental Health had lower than average participation rates for both English and Mathematics while

clients of the Addison and Rutland County programs had higher than average participation on both tests. Two other regional programs (Northwestern and Southeastern Vermont) had higher than average participation on the Mathematics tests but were not different for English test participation.

School Performance

Statewide, recipients of community mental health services were less likely than other students to score at or above standard on Vermont's standardized school tests. Only 42% of the service recipients scored at or above standard on the Mathematics test, compared to 64% of other students, and 49% of service recipients scored at or above standard on the English test compared to 65% of other students. There were no statistically significant differences between age or gender groups in Mathematics test performance among service recipients or in the general population. Females in the general population scored higher than males on the reading test, but there was no gender related difference in the service recipient population. Performance on the reading assessment decreased with increasing age in the general population and among service recipients, although the latter decrease was not statistically significant.

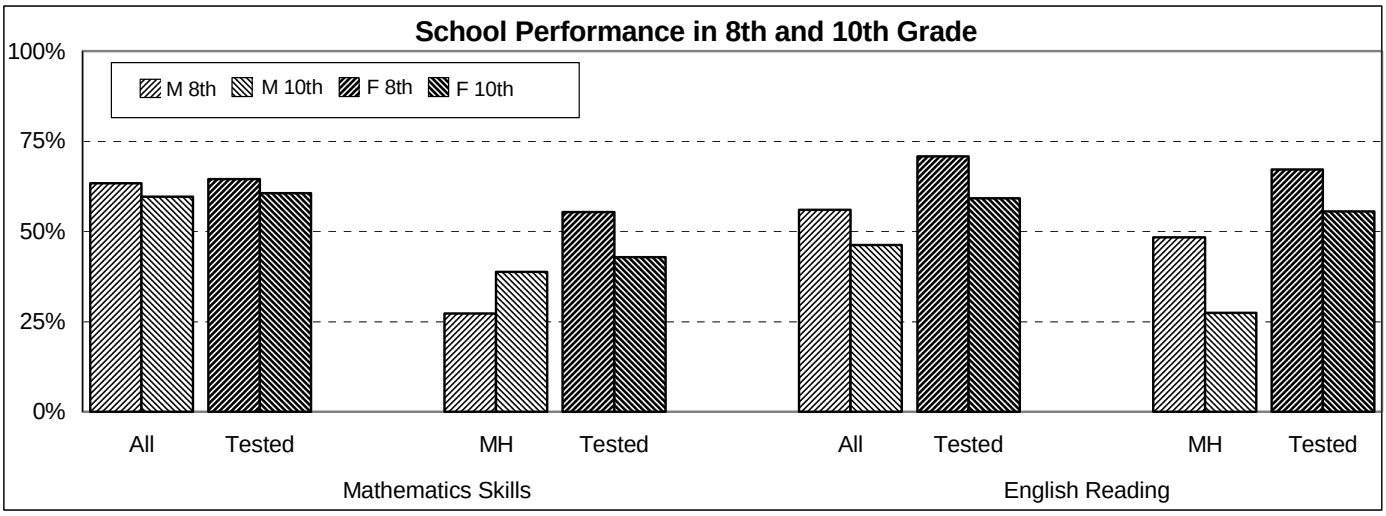
At the same time, however, service recipients' performance on the standardized tests was similar to the performance of other students at some local programs. In two regions of the state, test performance of service recipients was not different from other students on either the Mathematics or the Reading test. In four other regions of the state, the Reading test performance of service recipients was not different from other students, but the Mathematics test performance was lower.



Change in School Performance

For the student population as a whole, there was a slight decrease in school performance between 8th and 10th grade on the Mathematics test (64% to 62% at or above standard), but there was no change at all in the performance of students who had received mental health services during their 9th grade year. During this same period, there was a substantial decrease in performance on the Reading test (63% to 53%) for the student population as a whole, and a similar decrease in the performance of the students who had received mental health services (59% to 47%).

There are, however, substantial differences in the rate and direction of change for boys and girls who received mental health services. On the Mathematics test, the performance of boys increased (from 27% to 39% at or above standard) while the performance of girls decreased (from 55% to 43%). The cumulative effect of these changes, however, resulted in similar 10th grade performance on the Mathematics test for boys and girls who had received services. Performance on the English test, decreased between 8th and 10th grade for both boys and girls in the treatment group, but the decrease for boys (48% to 27%) was much greater than the decrease for girls (from 67% to 56%).



There were substantial differences among regions of the state in the amount and direction of change in school performance.

Discussion

We believe this pilot study has demonstrated the value of educational test scores for understanding and evaluating systems of care for children and adolescents. This rich new source of data can provide valid and reliable information on levels of school participation and school performance for recipients of mental health and other children's services programs. We believe two substantive findings of this project deserve special attention. First, the fact that gender differences in the school performance of service recipients are different than those for other students suggests the need for careful attention to the relationship between treatment variables, gender, and treatment outcomes. Second, the finding that elevated risk of trouble with the law decreases as time passes is encouraging. A similar pattern was found in our earlier three year prospective study of children's outcomes [4]. These results underline the importance of longer-term outcome studies.

The use of educational test scores that was demonstrated here should also be applied to other service sectors and geographical areas. Other service sectors should include child protection, juvenile justice, and special education programs. Other geographical areas should include urban areas and regions with race/ethnically diverse student populations. This research should consider the impact of both community characteristics (socio-economic, population density and diversity, etc) and individual student characteristics (both clinical and demographic) on levels of school participation and performance as well as the impact of community mental health and other services on school participation and performance.

Finally, future research should consider the impact of treatment modalities and service system characteristics on levels of school participation and performance. Differences between emerging evidence-based practice and other modalities should be explored. The impact of medication on school participation and performance is a particularly important area for investigation. In systems of care, the impact of service system characteristics such as caseload integration and service coordination on the relative school participation and performance of service recipients should be the focus of large-scale cross-region research.

Fortunately, the proliferation of electronic databases in conjunction with statistical technologies such as Probabilistic Population Estimation provide the opportunity for economical and effective research in all of the areas discussed above. This combination of data and methodology provides for the exploration of relationships between treatment variables and a variety of important treatment outcomes while protecting the confidentiality of medical records and the personal privacy of individuals [4]**.** This methodology is particularly valuable for examining longer term outcomes that go beyond the more traditional focus on

negative outcomes such as hospitalization and incarceration to include positive indicators such as employment and participation in post secondary education.

References

[1] Rosenblatt, A. (1998). Assessing the child and family outcomes of systems of care for youth with serious emotional disturbance. In M.H. Epstein, K. Kutash, & A. Duchnowski (Eds.) Outcomes for Children and Youth with Severe Emotional Disorders and Their Families, pp. 329-361. Austin, TX: Pro-Ed.

[2] National Association of State Mental Health Programs Directors. (1998) *Performance Measures of Mental Health Systems.* Arlington, VA. National Association of State Mental Health Program Directors. http://www.rdmc.org/nri/firstpage.htm

[3] Banks SM, and Pandiani JA (2001) Probabilistic Population Estimation of the Size and Overlap of Data Sets Based on Date of Birth. *Statistics in Medicine*, Vol. 20: 1421-1430.

[4] Pandiani, J.A., Schacht, L.M., Banks, S.M., (2001) After children's services: A longitudinal study of significant life events. *Journal of Emotional and Behavioral Disorders*, 9 (2): 131-138.

METHODOLOGICAL NOTE

PROBABILISTIC POPULATION ESTIMATION

Probabilistic Population Estimation is a statistical procedure that determines the number of people (with known confidence intervals) who are represented in data sets that do not contain unique person identifiers. Probabilistic Population Estimation uses information on the distribution of birth dates in a data set to determine the number of people represented in the data set. The number of people necessary to produce the number of birthdays observed in a single birth year cohort, for instance, would be calculated using the following formula:

$$P_j(l_j) = \sum_{i=1}^{l} \frac{365}{365-i}$$

where "P_j" is the number of people and "i" is the number of birth dates observed. Similar logic is used to determine the number of people who appear in more than one data set. The table below provides illustrative results of Probabilistic Population Estimation for populations of specified size.

Population Estimates for Specified Numbers of Birth Dates Within a Year

Birth Dates	Number of People	Birth Dates	Number of People
1	1.003 ± .103	180	249 ± 20
10	10.15 ± .776	250	423 ± 38
20	20.6 ± 1.54	300	632 ± 64
50	54. ± 4	330	860 ± 101
100	117. ± 9	360	1630 ± 325

POPULATION OVERLAP

In order to probabilistically determine the number of people shared across data sets that do not include a common person identifier, the sizes of three populations are determined and the results are compared. The number of people in each of the original data sets are the first two populations. The number of people in a data set that is formed by combining the two original data sets is the third data set.

The number of people who are shared by the two data sets is the difference between the sum of the numbers of people represented in the two original data sets and the number of people represented in the combined data set. This occurs because the sum of the number of people represented in the two original data sets includes a double count of every person who is represented in both data sets. The number of people represented in the combined data set does not include this duplication. The difference between these two numbers is the size of the duplication between the two original data sets, the size of the caseload overlap. In terms of mathematical set theory, the intersection of two sets is the difference between the sum of the sizes of the two sets (A+B) and the union of the two sets ($A \cup B$):

$$(A \cap B) = (A + B) - (A \cup B).$$

Related Reading

A Global Measure of Access to Mental Health Services for a Managed Care Environment. *Journal of Mental Health Administration*, Summer 1997. (Pandiani, Banks, & Gauvin)

A Methodology for Probabilistically Estimating Caseload Size and Overlap. The Evaluation Center @HSRI, January 1999 (Banks & Pandiani)

A Risk Adjusted Measure of Hospitalization Rates for Evaluating Community Mental Health Program Performance*. Administration and Policy in Mental Health*, March, 1999. (Pandiani, Banks, Schacht, & Gauvin)

After Children's Services: A Longitudinal Study of Significant Life Events. *Journal of Emotional and Behavioral Disorders*, 2001 Vol. 9 #2 (Banks, Pandiani, & Schacht)

Age and Mortality Among Problem Drinkers. *Addiction*. August 2000 (Banks, Pandiani, Schacht, & Gauvin)

Approaches to Risk Adjusting Outcome Measures Applied to Criminal Justice Involvement after Community Service. *Journal of Behavioral Health Services and Research*. 2001 Vol. 28 #3 (Banks, Pandiani, & Bramley)

Bed Closures and Incarceration Among Users of VA Behavioral Health Services in Upstate New York. *Mental Health Services and Research*, October 2000 (Rosenheck, Banks, Pandiani, & Hoff)

Caseload Segregation/Integration: A Measure of Shared Responsibility for Children and Adolescents. *Journal of Emotional and Behavioral Disorders*, Summer 1999 (Banks, Pandiani, & Schacht)

Consumer Satisfaction and Treatment Outcomes: Dissatisfaction with Mental Health Services and Incarceration after Treatment. *Administration and Policies in Mental Health*, 2001, Vol. 29, #2. (Pandiani, Schacht, & Banks)

Does Closing Inpatient Beds in One Public Mental Health System Result in Increased Use of Inpatient Services in Other Systems? *Psychiatric Services*, 2000 Vol. 2 #4. (Rosenheck, Banks, & Pandiani)

Elevated Risk of Being Charged with a Crime for People with a Severe and Persistent Mental Illness. *Justice Research and Policy,* Fall 2000 (Pandiani, Banks, Clements, & Schacht)

Personal Privacy vs. Public Accountability: A Technological Solution to an Ethical Dilemma. *Journal of Behavioral Health Services and Research*, November 1998. (Pandiani, Banks, & Schacht)

Practice Patterns and Hospitalization Rates: A Statewide Program Evaluation. *Administration and Policy in Mental Health*, September, 1998. (Banks, Pandiani, Gauvin, Reardon, Schacht, & Zovistoski)

Probabilistic Population Estimation of the Size and Overlap of Data Sets Based on Date of Birth. *Statistics in Medicine*, 2000. (Banks & Pandiani).

The Use of State and General Hospitals for Inpatient Psychiatric Care. *American Journal of Public Health*, March 1998. (Banks & Pandiani)

Utilization of Local Jails and General Hospitals by State Psychiatric Center Patients. *The Journal of Behavioral Health Services and Research*, November 2000 (Banks, Stone, Pandiani, Cox, & Morchauser)

Using Existing Data Bases to Measure Treatment Outcomes. In *Developing Outcome Strategies in Children's Mental Health*. Edited by Hernandez, M and Hodges, S. Paul H. Brookes Publishing Co, Baltimore 2001 (Banks & Pandiani)

Using Incarceration Rates to Measure Mental Health Program Performance. *Journal of Behavioral Health Services and Research*, August 1999. (Pandiani, Banks, & Schacht)

For more examples of VT PIP research visit: www.state.vt.us/dmh/Data/PIPs/pipindexchildren.htm

For more information on Probabilistic Population Estimation visit: www.thebristolobservatory.com